EPA Document # 815-R-05-006

Statistical Protocol for the Determination of the Single-Laboratory Lowest Concentration Minimum Reporting Level (LCMRL) and Validation of Laboratory Performance at or Below the Minimum Reporting Level (MRL)

Prepared for:

Mr. David Munch
and
Ms. Phyllis Branson
U.S. Environmental Protection Agency
Office of Ground Water and Drinking Water
Standards and Risk Management Division
Technical Support Center
26 W. Martin Luther King Drive (MS-140)
Cincinnati, OH 45268

November 2004

Acknowledgments

This document has been prepared in conjunction with the Environmental Protection Agency's (USEPA) Office of Ground Water and Drinking Water, Standards and Risk Management Division, Technical Support Center (TSC). David Munch served as USEPA TSC's Work Assignment Manager. Chris Frebis of USEPA's TSC, Elizabeth Hedrick of EPA's National Exposure Research Laboratory/Office of Research and Development, and Stephen Winslow and Barry Pepich of Shaw Environmental and Infrastructure, Inc. provided technical guidance. The Cadmus Group, Inc. served as primary contractor. Dr. George Hallberg served as Project Manager, and the document was prepared by John Martin. This document is the product of more than one year of interactive input from the above named individuals, peer reviewers, and laboratory personnel who participated in a proof-of-concept interlaboratory study from November 2003-March 2004. The input of the following individuals is acknowledged:

Richard H. Albert, Ph.D., United States Food and Drug Administration (retired) - peer review of initial documents used to design the interlaboratory study and in the determination of Lowest Concentration Minimum Reporting Levels/Minimum Reporting Levels (LCMRLs/MRLs);

William Horwitz, Ph.D., AOAC International - peer review of initial documents used to design the interlaboratory study and in the determination of LCMRLs/MRLs;

William T. Foreman, Ph.D., United States Geological Survey - peer review of initial documents used to design the interlaboratory study and in the determination of LCMRLs/MRLs;

Andrew D. Eaton, Ph.D., MWH Laboratories - peer review of initial documents used to design the interlaboratory study and in the determination of LCMRLs/MRLs and laboratory analyses;

William Telliard and personnel from EPA's Engineering and Analysis Division - peer review of initial documents used to design the interlaboratory study and in the determination of LCMRLs/MRLs;

Brad Venner of USEPA's National Enforcement Investigations Center - peer review of initial documents used in the determination of LCMRLs/MRLs;

Michael D. Wichman, Ph.D., University Hygienic Laboratory - laboratory analyses;

Phillip Godorov, Philadelphia Water Department Laboratory - laboratory analyses;

John V. Morris, EPA Region 5 Laboratory - laboratory analyses; and

Lisa Wool, EPA Region 6 Laboratory - laboratory analyses.

Mention of trade names and/or laboratories does not constitute endorsement or recommendation for use.

Contents

Exhibits

List of Acronyms

AML	Alternate Minimum Level
ASTM	American Society for Testing Materials
CFR	Code of Federal Regulations
HR_{PIR}	Half-Range Prediction Interval of Results
ICR	Information Collection Rule
IDC	Initial Demonstration of Capability
IDL	Instrument Detection Level
IQCL	Instrument Quality Control Level
$IQE_{Z\%}$	Interlaboratory Quantitation Estimate (with Z% Relative Standard Deviation)
LCMRL	(Single laboratory) Lowest Concentration Minimum Reporting Level
LRL	Laboratory Reporting Level
LT-MDL	Long Term Method Detection Limit
MCL	Maximum Contaminant Level
MDL	Method Detection Limit
ML	Minimum Level
MQCL	Method Quality Control Level
MRL	Minimum Reporting Level
NWQL	National Water Quality Laboratory
OGWDW	Office of Ground Water and Drinking Water
OLS	Ordinary Least Squares
PE(S)	Performance Evaluation (Study)
PIR	Prediction Interval of Results
PQL	Practical Quantitation Limit
PWS	Public Water Systems
QCL	Quality Control Level
RSD	Relative Standard Deviation
UCMR	Unregulated Contaminant Monitoring Regulation
USEPA	United States Environmental Protection Agency
USGS	United States Geological Survey
VWLS	Variance Weighted Least Squares

1.0 INTRODUCTION

The Safe Drinking Water Act Amendments of 1996 require EPA to establish criteria for a monitoring program and to publish a list of not more than 30 unregulated contaminants for which public water systems (PWS) are to monitor. The monitoring program will provide a national assessment of the occurrence of these contaminants in public drinking water, that will be used to help decide which contaminants may or may not require regulation in the future. In 1999, EPA revised the approach for unregulated contaminant monitoring in the Unregulated Contaminant Monitoring Regulation (UCMR) (64 FR 50556; USEPA, 1999) and subsequent revisions.

PWSs will be required to monitor for a variety of contaminants under the UCMR. A Minimum Reporting Level (MRL) will be assigned to each contaminant, and laboratories will be required to report all occurrences of these contaminants at concentrations that are equal to or greater than the established MRL. The MRL has been developed, in part, as an alternative to the Practical Quantitation Limit (PQL) which, in the past, has been determined by either evaluating EPA Performance Evaluation Study (PE) data or by applying a multiplication factor to the Method Detection Limit (MDL) (as described at 40 CFR Appendix B to Part 136), that had been established during the development of the analytical method. The MRL differs from the MDL by considering not only the standard deviation of low concentration analyses (precision), but also the accuracy of the measurements. In addition, since the privatization of PE programs by EPA, the data that are required for PQL determination are no longer readily available. The MRL was introduced with the new analytes and new methods for implementing the new UCMR. The MRL may be useful as an alternative to the PQL for setting future regulatory limits, as well.

MRLs have often been determined by analytical laboratories using expert professional judgement, but consistent criteria for MRL determination have not been established. In both the Information Collection Rule (ICR) and the UCMR, OGWDW specified MRLs and an accuracy requirement for recovery at the MRL so that data quality was documented daily. The most difficult issue for the MRL has been developing a consistent procedure to set the MRL. EPA has developed a statistical approach for determination of single-laboratory Lowest Concentration MRLs (LCMRLs) using linear regression and prediction intervals. This approach has been evaluated through expert review and through the performance of a pilot-scale interlaboratory study.

The MRL is the lowest analyte concentration which demonstrates known quantitative quality. The LCMRL, as calculated by the procedure presented in this paper, is the lowest true concentration for which the future recovery is predicted to fall, with high confidence (99%), between 50 and 150% recovery. A result below the MRL is considered to be an estimated value that does not satisfy quality control objectives. It should be noted that the decision to report estimated data is dependent upon the objectives of a study and not a point of discussion here.

In this paper, we present a systematic procedure for determining an LCMRL. The LCMRL is used to determine the MRL for an analyte by using either a multiplying factor or by

pooling the results from a multi-laboratory study. Drinking water laboratories will confirm that they are capable of meeting a required MRL during their Initial Demonstration of Capability (IDC) using the procedure described in Section 6.0 of this paper. The LCMRL protocol is based on a linear regression procedure as described in Section 2.0; the statistical basis for the LCMRL is described in Section 3.0; possible procedures for the determination of MRLs from LCMRLs are described in Section 4.0; example LCMRL determinations are presented in Section 5.0; the statistical basis of MRL validation is presented in Section 7.0; and an example of validation of laboratory performance at or below the MRL is presented in Section 8.0. It should be noted that three distinct procedures are presented in this document:

- LCMRL - determined by selected laboratories during method development. All laboratories are encouraged to determine LCMRLs that are unique to their laboratory, as this may aid them in establishing spiking concentrations for validation of performance at or below an MRL; however, this is not required.
- Validation of laboratory performance at or below an MRL - performed by each laboratory that is analyzing samples for UCMR as part of IDC.
- Daily check at or below an MRL - performed daily by each participating laboratory to demonstrate meeting Data Quality Objectives (DQOs) for all UCMR analytes (DQO = within 50 - 150% recovery).

2.0 DETERMINATION OF LCMRL DURING METHOD DEVELOPMENT

This section includes a description of the steps in the process used to determine the LCMRL for a particular analyte and several key procedural issues. This step is limited to laboratories that are participating in method development and to those that desire to establish laboratory-specific LCMRLs.

2.1 Determination of the LCMRL

The range of instrument calibration standards must encompass the levels being evaluated, otherwise the determined LCMRL may be unreliable. The instrument calibration range and the type of regression model used to fit the instrument calibration data for the LCMRL determination must be used in the future as the normal day-to-day calibration for the analyte and method in question. The calculated LCMRL cannot be lower than the lowest calibration standard.

It is preferable that, at each of at least four levels of determination, a minimum of seven replicate samples in reagent water are processed through the entire method procedure (including extraction and all preservatives, where applicable). At an absolute minimum, five samples at each of four concentrations, or seven samples at each of three concentrations are processed through the entire method procedure. An initial estimate of spiking level should consider a concentration of reliable quantitation and an analyte peak area at least three times greater than

that found in a blank sample processed through the entire method procedure.

The LCMRL is determined via the following five steps, using EPA's Internet-accessible LCMRL calculator or commercially available statistical software with the requisite capabilities. The following steps are performed by the LCMRL calculator, but are presented here for transparency and to allow for the optional use of other statistical software.

- For each analyte, the spiked concentrations (x-axis) are plotted against the measured concentrations (y-axis);

- The LCMRL data are regressed using Ordinary Least Squares (OLS), with a straight line regression equation and a 99% prediction interval around the regression line. Do not force the regression line through the origin. A test for constant variance over the range of spiking concentrations must be performed. The threshold for passing the test of constant variance is 1%. This corresponds to a probability of 0.01 for concluding that the variance is not constant with respect to concentration when it actually is constant. If the data do not pass the test of constant variance, a Variance Weighted Least Squares (VWLS) regression model must be employed. Details regarding the test for constant variance and the VWLS model are presented in Section 2.2.1 of this document.

- Plot or draw lines that correspond to 150% and 50% recovery of the spiked concentration.

- Drop a perpendicular to the x-axis starting from the point at which the upper prediction interval line intersects the 150% recovery line. Drop a second perpendicular to the x-axis starting from the point at which the lower prediction interval line intersects the 50% recovery line. The location of the perpendiculars and the LCMRL are determined mathematically as follows:

  - Assume that the 99% prediction interval is a straight line between the two known true (spiked) concentrations that encompass the intersection in question;

  - Determine the slope (m) of the prediction interval between the two true concentrations ($m = \Delta y/\Delta x$);

  - For a given x and y, use m to calculate the y-intercept, b;

  - For the upper recovery line, $y = 1.5x$; for the lower recovery line, $y = 0.5x$;

  - Since at the intersection of the prediction interval boundary and the recovery line, there is one value for y, $mx + b = 1.5\ x$ (for the intersection at the upper recovery line) and $mx + b = 0.5\ x$ (for the intersection at the lower recovery line). The LCMRL equals the larger value of x. If the prediction interval is contained

entirely within the 50 to 150% recovery range, the LCMRL is set equal to the lowest spiked concentration. In cases where the prediction interval is located entirely outside of the recovery range or is located outside the recovery range at the highest spiked concentration, the LCMRL is indeterminable, at a concentration that is greater than the highest spiking concentration.

- The LCMRL for a particular analyte is the larger of the two values indicated by the intersection of the perpendiculars with the x-axis; however, the LCMRL cannot be less than the lowest spiked concentration or lowest calibration standard for a particular analyte.

2.2 Diagnostic Procedures and Potential Issues

To avoid the perception of a "black box" model, and to maintain transparency of this statistical protocol, the instrument calibration regressions, the replicate analyses, and the regression data should be analyzed to determine whether the regression and associated prediction intervals have been appropriately modeled. For the purposes of this procedure, a test of constant variance is considered to be the most appropriate diagnostic procedure for determining the validity of the regression model.

2.2.1 Diagnostic Procedures

Prior to the analysis of the replicate samples, instrument calibration curves should be evaluated for goodness-of-fit to the calibration data. Hence, each calibration curve must be evaluated to determine whether the calibration curve data are linear. For some methods and/or analytes, a quadratic regression equation may provide the better fit for the instrument calibration data. The fit of the instrument calibration equation is indicated by the *adjusted* coefficient of determination (R^2) value. As the degree of the polynomial increases, the number of predictors also increases. This results in a larger R^2 value. While a larger R^2 value implies a better fit, the model with the larger *adjusted* R^2 value represents the better fit. The R^2 value explains the proportion of the variation in the dependent variable that is explained by the regression equation; the *adjusted* R^2 value accounts for the dependence of R^2 on the degrees of freedom. Hence, the *adjusted* R^2 value addresses relative variance, where variance equals variation divided by the degrees of freedom.

The adjusted coefficient of determination, $R^2_{adj} = 1 - \left[\frac{n-1}{n-p} (1 - R^2) \right]$

where: n = the number of observations; and
p = the number of predictors, or the number of independent variables in the regression equation (including the constant).

The best-fit regression equation may be necessary to produce linear results for the LCMRL data. However, most analytical instrumentation software calculates either the

correlation coefficient (R) or the R^2 value. If a value for R^2 of 0.99 (i.e., an R value of 0.995) or greater is obtained for the instrument calibration regression, the use of a better fit regression model will likely have little effect on the LCMRL that is obtained. To obtain the lowest value for the LCMRL, a linear polynomial regression equation must be employed during the analysis of the replicate samples as part of the LCMRL determination process.

While several methods for testing the assumption of constant variance are available, the Cook-Weisberg test is used in this procedure. A quantity, S, is computed by dividing the Mean Sum of Squares (MSS) by two. S has a χ^2 distribution, with one degree of freedom. The data pass the constant variance test if the probability (P) > χ^2 is 0.01 or greater. For a null hypothesis of "the variance is constant", this means that there is a 1% probability of Type I error, or concluding that the variance is not constant when it actually is constant. In the VWLS model, the measured and spiked concentration data are weighted by dividing each value by the standard deviation (i.e., the square root of the variance) of the measured concentrations at its corresponding spiking concentration. An important assumption of VWLS is that the population variance and standard deviation at each spiking concentration are known. The variance and standard deviation that are obtained from the LCMRL replicate data (i.e., sample variance and standard deviation) are used in this procedure. Another assumption inherent to this process is that the replicate data are normally distributed.

2.2.2 Potential Issues

- **Range of Spike Concentrations:** The range of spike concentrations must be contained within the range of instrument calibration standards, and this range must be used as the routine daily calibration range for subsequent analyses. The range of instrument calibration and spike concentrations should not exceed two orders of magnitude, except in relatively rare cases of extended linearity, such as certain metals analyses, where the analytical method and/or instrumentation specifies a broader range. If the magnitude of the difference between the upper and lower spike concentrations is greater than approximately 10-20 times, adjustment of the linear scale range may be necessary to obtain the resolution needed to visually determine LCMRLs from the intersection of the perpendicular with the x-axis.

- **Linearity and Non-Linearity of Analytical Data:** The replicate analyses may result in a non-linear relationship between the spiked and measured concentrations for certain analytes and/or analytical methods. For an analyte and/or analytical method, the following are some of the factors that may result in non-linearity of the LCMRL regression curve:

    - the better-fit model (i.e., straight line or quadratic) was not utilized for the instrument calibration curve; and
    - the replicate analyses were performed using more than one calibration curve; and

As previously mentioned, it is useful to evaluate the fit of the instrument calibration curves for each analyte prior to the analysis of the replicate samples so the LCMRL data are as close to linear as possible. Linear and quadratic regression models should be applied to the instrument calibration data, the *adjusted* R^2 values should be compared, and the model with the largest *adjusted* R^2 values should be selected as the best fit calibration model. However, if a value for R^2 of 0.99 or greater (i.e., $R \geq 0.995$) is obtained, the use of a better fit instrument calibration regression model will likely have little effect on the resultant LCMRL.

- **Outliers in the Data Set:** The effect of a potential outlier on the determination of the LCMRL depends on the magnitude of the residual error and the distance of the outlier from the mean of the spiked data. While outliers may represent actual laboratory conditions, the presence of outliers may result in artificially high or even indeterminable LCMRLs. If the reason for an outlier is known and justified (e.g., analyst error), or if the outlier is identified with specified (99%) confidence based on the results of Dixon's Q Test (described below), an outlier may be omitted from the LCMRL determination on a case-by-case basis. Note that the outlier exclusion process for LCMRL determination is limited to a single outlier at a single concentration. Outliers may not be excluded from data sets used to demonstrate validation of laboratory performance at or below an MRL.

While there are several tests that can be performed to evaluate outliers, Dixon's Q Test has been widely proposed for application in analytical chemistry evaluations (Rorabacher, 1991) and is relatively simple to apply. The test for detecting a single outlier in a data set is derived from Rorabacher (1991) in Equation 1. The original terminology presented in the paper has been modified somewhat for clarification:

$$(1) \qquad Q = \frac{|x_o - x_c|}{x_{hi} - x_{lo}}$$

where: x_o = the potential outlier;
x_c = the closest value to the potential outlier;
x_{hi} = the highest value in the data set (including the outlier where applicable); and
x_{lo} = the lowest value in the data set (including the outlier where applicable).

The calculated value of Q is compared to tabulated critical values of Q. If the calculated Q exceeds the critical value of Q, the outlier is rejected at the confidence level from which the critical value was taken. In this test, a confidence level of $\alpha = 99\%$ was used. This corresponds to rejecting only 1% of values as outliers when they are not truly outliers (Type I error). For the testing of a single outlier, Q values for the r_{10} test, as presented in Table 1 of Rorabacher (1991) were used. The designation "r_{10}" refers to 1 outlier on one "end" of a data set, and 0 outliers on the opposite "end" of the data set. The critical values of Q for n = 5 to 30 are presented in

Exhibit 1 (Rorabacher, 1991).

Exhibit 1: Critical Values of Q at the 99 % Confidence Level

Replicates	Critical Value of Q	Replicates	Critical Value of Q
5	0.821	18	0.442
6	0.74	19	0.433
7	0.68	20	0.425
8	0.634	21	0.418
9	0.598	22	0.411
10	0.568	23	0.404
11	0.542	24	0.399
12	0.522	25	0.393
13	0.503	26	0.388
14	0.488	27	0.384
15	0.475	28	0.38
16	0.463	29	0.376
17	0.452	30	0.372

These critical values of Q are presented to allow for analysis of outliers for a variety of LCMRL evaluations, from a single laboratory analyzing a minimum of five replicates, to a pooled evaluation involving three laboratories, each analyzing ten replicates at the concentration in question.

3.0 BACKGROUND AND STATISTICAL BASIS OF THE LCMRL

Confidence in quantitation depends on measurement accuracy as defined by precision and bias. The determination of the quantitation limit has a lengthy history. Many past procedures for quantitation limit determination used multiples of sample standard deviation of blank signals or at low-level fortification, but did not consider the bias of measurement.

3.1 History of Selected Detection and Quantitation Procedures

International Union of Pure and Applied Chemistry (IUPAC) (Currie) Detection Limit Procedure

The Currie detection limit procedure (Currie, 1968; Currie, 1999) describes three types of detection limit relations:

Critical level (L_C). The critical level, L_C, is the lowest value that, with specified confidence, does not result from a blank. The probability of exceeding L_C when analyte is absent is α. A value for α of 0.01 signifies the interval at or above L_C should contain only 1% false positives. The L_C is a minimum value of estimated net signal or concentration applied against background noise.

Detection limit (L_D). The detection limit (L_D) is the minimum detectable value of the net signal (or concentration) for which the false negative error is β, which is the probability that a true value at the L_D is not measured as less than or equal to L_C. Given a normal distribution of results, when samples contain an analyte at the L_D, there is a 50% chance that analyzed results will fall below this limit and not be reported (i.e., a false negative).

Determination or Quantitation limit (L_Q). The quantitation limit (L_Q) marks "the ability of the chemical measurement process to adequately 'quantify' an analyte." Replicate analysis at L_Q will produce estimates with a relative standard deviation ($\%RSD_Q$), such as the 10% RSD mentioned by Currie.

Currie (IUPAC) procedure issues. Two issues with the IUPAC procedure are the lack of bias accountability for the quantitation limit, and the difficulty with the determination of blank variance in chromatographic methods. Since variance from replicate blanks determines the region of reliable quantitation, there is not an accuracy requirement for the quantitation limit. Measurement bias at low level is not addressed except to say that the bias bounds "require skilled and exhaustive scientific evaluation of the entire structure of the chemical measurement process."

The EPA Method Detection Limit

The EPA's Method Detection Limit (MDL) procedure (40 CFR 136, Appendix B) avoids the problems of determining variance at zero concentration by fortifying samples at low levels which must be 1 to 5 times the calculated estimated MDL. The MDL is defined as:

(2) $MDL = t_{(n-1,1-\alpha)} * s$

where: t = Student's t;
s = the standard deviation of replicate spikes at low-level;

1-α = the probability point; and
n-1= degrees of freedom.

The derivation is found in Glaser *et al.* (1981). The USEPA's MDL procedure uses the standard deviation from low-level fortified replicates to estimate a confidence interval around zero concentration that includes 99% of all false positives.

Standard Methods (17th Edition, 1989) Method Detection Procedure

The Standard Methods detection level procedure (APHA, AWWA, WPCF, 1989) uses terms common to Currie's and USEPA MDL procedures. Standard Methods describes the instrument detection level (IDL) as "the constituent concentration that is at least five times the signal-to-noise ratio of the instrument." The IDL is determined using non-extracted standards. Standard Methods refers to this as a "critical level," but is quite different from Currie's critical level.

Standard Methods defines the lower level of detection (LLD) as "the constituent concentration in reagent water that produces a signal 2 x (1.645) x *s* above the mean of blank analyses" (i.e., twice the IDL). This sets both Type I error (α is the rate of false positives) and Type II errors (β is the rate of false negatives) at 5%.

Standard Methods describes the method detection limit (MDL) in a manner similar to the EPA MDL. Standard Methods qualifies the use of its MDL, by saying it "can be achieved by experienced analysts operating well-calibrated instruments on a non-routine basis."

The level of quantitation (LOQ) is defined as "the constituent concentration that produces a signal sufficiently greater than the blank that it can be detected within specified levels by good laboratories during routine operating conditions." Typical concentration is "10 *s* above the reagent water blank signal."

Long-Term Minimum Detection Level (LT-MDL) of the USGS NWQL

The United States Geological Survey National Water Quality Laboratory (NWQL) has begun to use a reporting procedure based on a Long-Term Method Detection Level (LT-MDL) (Childress et al., 1999). For the LT-MDL concentration, the risk of a false positive is set to no more than 1 percent. At the LT-MDL concentration though, the risk of false negative is up to 50%. A laboratory reporting level (LRL) is set at twice the LT-MDL and concentrations measured between the LRL and the LT-MDL are reported as estimated concentrations. The data user is given the flexibility to censor the estimated data. Non-detections are censored at the LRL. The LT-MDL is determined over an extended time by using all method instrumentation and large number of replicate spike samples to obtain a more accurate and realistic measurement of the standard deviation near the MDL.

Alternate Minimum Level (AML)

Gibbons, Coleman, and Maddalone suggested in 1997 the use of an alternate minimum level (AML) which uses a multiple-concentration calibration procedure (Gibbons et al. 1997). They felt that the standard deviation, *s*, used to calculate the minimum level (ML) was faulty because *s* depended on the choice of spiking concentration due to non-constant variance. The AML takes into account the uncertainty in the calibration function and in the standard deviation.

The determination requires spiked samples at levels below the initially estimated MDL value to many times the estimated MRL. The determination of the AML involves 6 steps which requires 11 calculations including an exponential regression model for standard deviation versus concentration. While this is a very interesting procedure, the math may be too complex for routine application. The AML calculations are sold as a software package by St@tServ - Statistical Software at: http://www.statserv.com/softwares.html. It should be noted that the AML is a detection-related value and not a quantitation-related value.

ASTM International's Interlaboratory Quantitation Estimate ($IQE_{Z\%}$)

The following description is "adapted from ASTM D 6512-00 Standard Practice for Interlaboratory Quantitation Estimate, copyright ASTM International, 100 Barr Harbor Drive, West Conshohocken, PA 19083. The information is used with permission; ASTM International, however, is not responsible for any changes made by the Exchange."

The $IQE_{Z\%}$ is the lowest concentration for which a single measurement from a laboratory selected from the population of qualified laboratories will have an estimated Z% RSD (relative standard deviation), where Z is dictated by data quality objectives. This procedure uses a regression approach to determine the point of 10% RSD among cross-lab mean values, with no simplifying assumptions about the dependence of standard deviation on concentration. The IQE is a minimum concentration at which most laboratories can be expected to reliably measure a specific chemical contaminant during day-to-day analyses. The procedure is an interlaboratory extension of the RSD approach used in the Gibbons AML procedure. In addition, the $IQE_{Z\%}$ basically corresponds to the L_Q (Currie, 1968), as the lowest concentration that produces Z% RSD.

Quality Control Level (QCL)

The QCL was introduced in 1994 as a quantity that "determines the lowest concentration that meets the data quality objectives of the data user in terms of the minimum acceptable precision and accuracy" (Kimbrough and Wakakuwa, 1994). A Method Quality Control Level (MQCL) is determined by first determining an Instrument Quality Control Level (IQCL) (i.e., interference-free) based on user-specified bias and RSD criteria. These same bias and RSD criteria are then applied to matrix and method-specific conditions to determine the MQCL for a given analyte, matrix, and analytical method. One estimate of the MQCL is determined from the

IQCL and any correction factors related to extraction, concentration, digestion, etc. A second estimate is obtained by spiking aliquots of the matrix with analytes to generate solutions of concentration that are equal to, less than, and greater than the first estimate of the MQCL. The lowest concentration that meets the bias and RSD criteria is the second estimate of the MQCL. A check is then performed at the second estimate of the MQCL to demonstrate whether the bias and RSD criteria are met.

3.2 Basis of the LCMRL

The existence of several methods for establishing detection and quantitation levels has created a need for uniformity in the process. EPA considered the procedures described in Section 3.1, as well as others, and decided that a regression/prediction interval approach that also combines desirable features of these procedures, with consideration for ease of application, transparency, and cost, would best meet the objectives of the UCMR. Thus, the MRL described in this paper is proposed as a quantitation metric that considers not only the standard deviation of low concentration analyses (precision), but also the bias of the measurements. The predefined QC interval and the confidence level of the Student's *t* value are quality assurance objectives that can be tailored to fit future analytical and policy needs. The decision on how, or if, to report values below the MRL will depend on the objectives of the study being conducted. The QC interval of recovery chosen for use in this paper, 50 to 150%, is based upon experienced judgement from chemical analysts. The prediction interval for the regression line that is derived from the Student's *t* distribution was chosen as 99% because it is conservative, consistent with other DQO's used in this procedure, it minimizes false positives, and is often used in other statistical tests. It should be noted that this procedure is designed for data that are continuous (e.g., Gaussian) rather than with data that are discrete, such as "counting" methods.

Ordinary Least Squares (OLS) is used to fit a regression of the measured concentrations against the spiked concentrations. In a simple linear regression, OLS draws a line through the data in a way that minimizes the sum of the squares of the deviations of the observed values from the regression line. As the regression line may be a higher order polynomial, the general form of the model is given by (StataCorp, 2001):

$$(3) \qquad y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \cdots + \beta_p x_i^p + e_i$$

where: y_i = the measured concentration for observation i;
x_i = the spiked concentration for the observation i;
p = the order or degree of the polynomial (and hence the number of predictors); and
e_i = the residual error term.

The standard error (SE) is the sum of the residual error and the error of the prediction. Given this standard error, a 99% prediction interval can be constructed around the regression line. The 99% prediction interval is given by:

$$(4) \quad \hat{y} \pm t_{(1-\alpha/2,\, n-p-1)} * SE$$

where: $\hat{y}$ = the estimated value of y;
$t_{(n-p-1)}$ = the value of the t distribution with (n-p-1) degrees of freedom that is exceeded with a probability of (1-0.99)/2, or 0.005; and
SE = the standard error.

The probability is (1-0.99)/2 because there is a 1 percent probability the true value is greater than or less than the predicted value; i.e., there is a 0.5% probability it is higher and a 0.5% probability it is lower. Thus, for a true concentration, x, a future measured concentration, y, is predicted to fall, with 99% confidence, within the prediction interval described by Equation 4. Additional statistical background regarding the LCMRL is presented in "Evaluation of the Lowest Concentration Minimum Reporting Level (LCMRL) and the Minimum Reporting Level (MRL): Primary Analyte Analysis," which is available in the EPA Docket.

4.0 DETERMINATION OF MRLs FROM LCMRLs

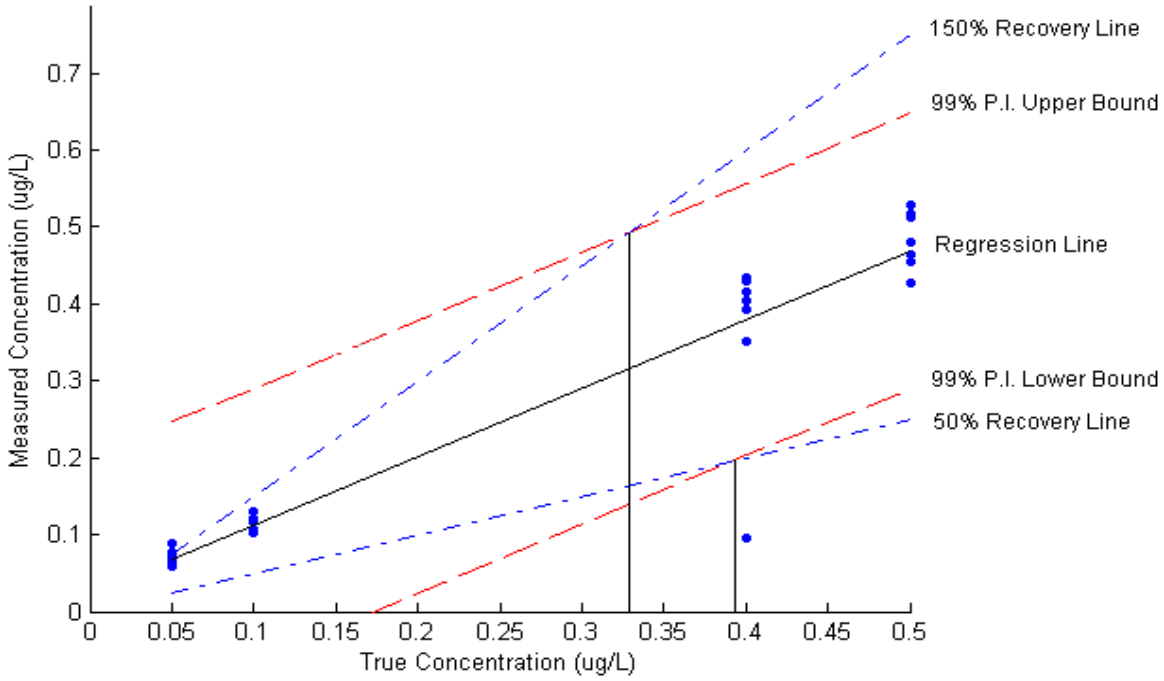
During analytical method development as part of the UCMR, laboratories determine their own values for the LCMRL. The mean of these LCMRL values was calculated for each analyte. In cases where data from three or more laboratories were available, three times the standard deviation of the LCMRLs was added to the mean of the LCMRLs. In cases where data from two laboratories were available, three times the difference of the LCMRLs was added to the mean of the LCMRLs. In statistical theory (Chebyshev's Inequality), three standard deviations around the mean incorporates the vast majority (at least 88.9%) of the data points. In the case where there are only two laboratories, the difference serves as a surrogate for the standard deviation, because of the uncertainty in the estimate of the standard deviation with only two data points. The MRL for each analyte was determined by rounding this number to two significant digits. Note that this procedure differs from that presented in the LCMRL/MRL Evaluation report (Primary Analyte Analysis), since the results of the primary analyte evaluation were used to decide how LCMRLs would ultimately be determined from multiple laboratory data.

5.0 EXAMPLE LCMRL DETERMINATIONS

Exhibits 2-5 demonstrate the use of OLS and VWLS with and without an outlying datum. As part of an interlaboratory study that was conducted to evaluate the LCMRL/MRL concept, replicate data for atrazine by EPA Method 507 were generated by analyzing samples fortified at

0.05, 0.1, 0.4, and 0.5 μg/L. Seven replicates were analyzed at each of these concentrations. Note that the results of the constant variance test for the OLS regression presented in Exhibit 2 indicate that variance is not constant, but changes with concentration. For illustrative purposes, the LCMRL is determined to be 0.39 μg/L; however, OLS is not appropriate for use in cases of non-constant variance. The appropriate use of VWLS for these data is presented in Exhibit 3. Note also that an outlying datum at a true (spiked) concentration of 0.4 μg/L is present in the data sets that are used in Exhibit 2 and Exhibit 3. Exclusion of the outlying datum and reevaluation of the LCMRL by OLS and VWLS is presented in Exhibit 4 and Exhibit 5, respectively. Criteria for determining the validity of outlier exclusion in the LCMRL determination process are presented in Section 2.2.2.

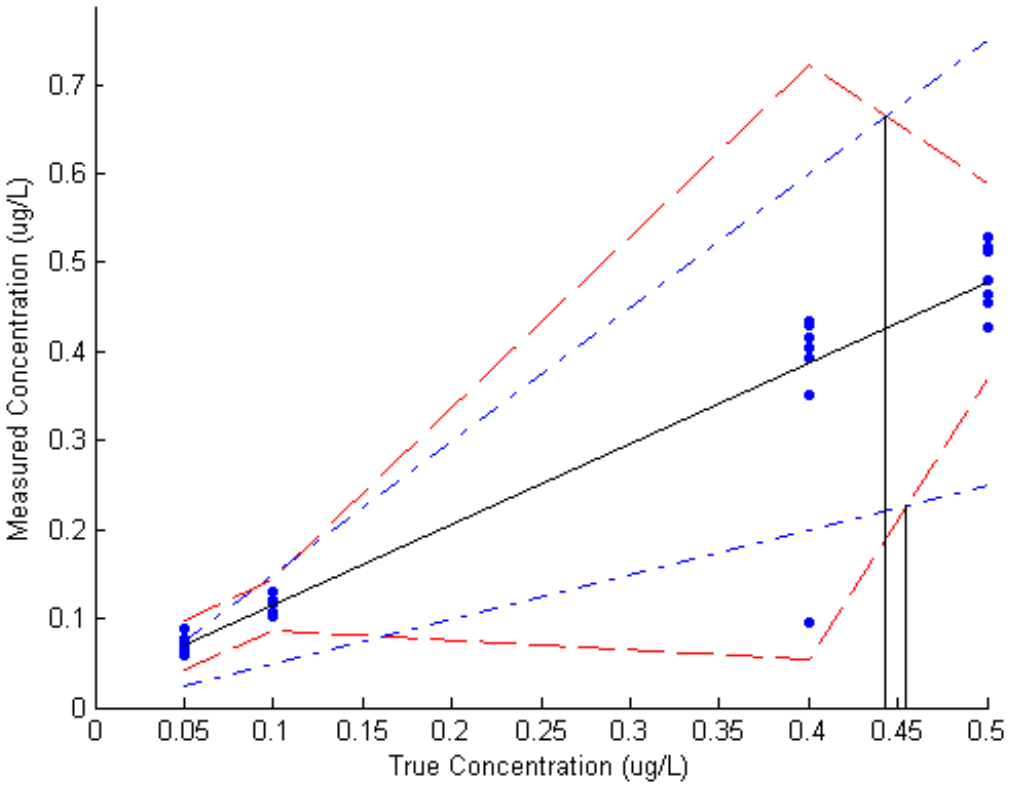
Exhibit 2: Atrazine by EPA Method 507
LCMRL Determination by Ordinary Least Squares
All Data



$R = 0.9439$ $R^2 = 0.8910$ Adj. $R^2 = 0.8868$
$y = 0.890(x) + 0.023$
Constant Variance Test: Failed ($P>Chi^2(1) = 0.0052$)
LCMRL = 0.39 µg/L

Since the variance is not constant, VWLS is used to regress the data as shown in Exhibit 3.

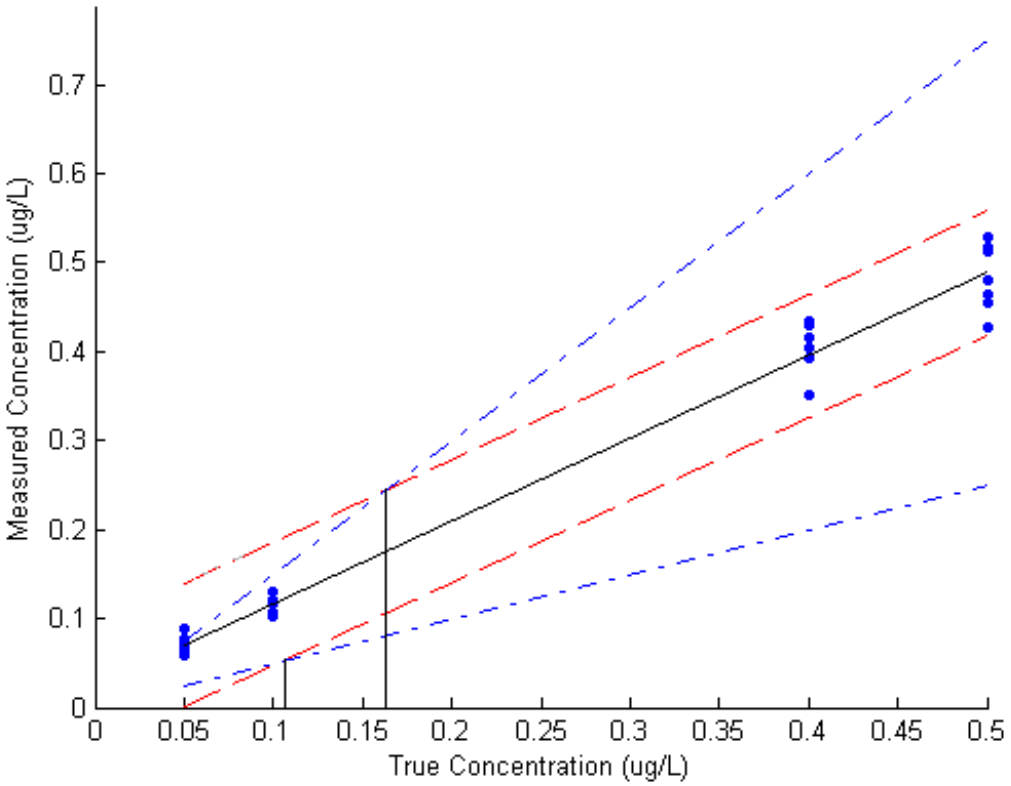
Exhibit 3: Atrazine by EPA Method 507
LCMRL Determination by Variance Weighted Least Squares
All Data



y = 0.909(x) + 0.024
LCMRL = 0.45 µg/L

The presence of the potentially outlying datum at a true (spiked) concentration of 0.4 µg/L has a significant effect on the prediction interval and the resultant LCMRL. Exclusion of the outlier allows for additional LCMRL determinations, as presented in Exhibit 4 and Exhibit 5.

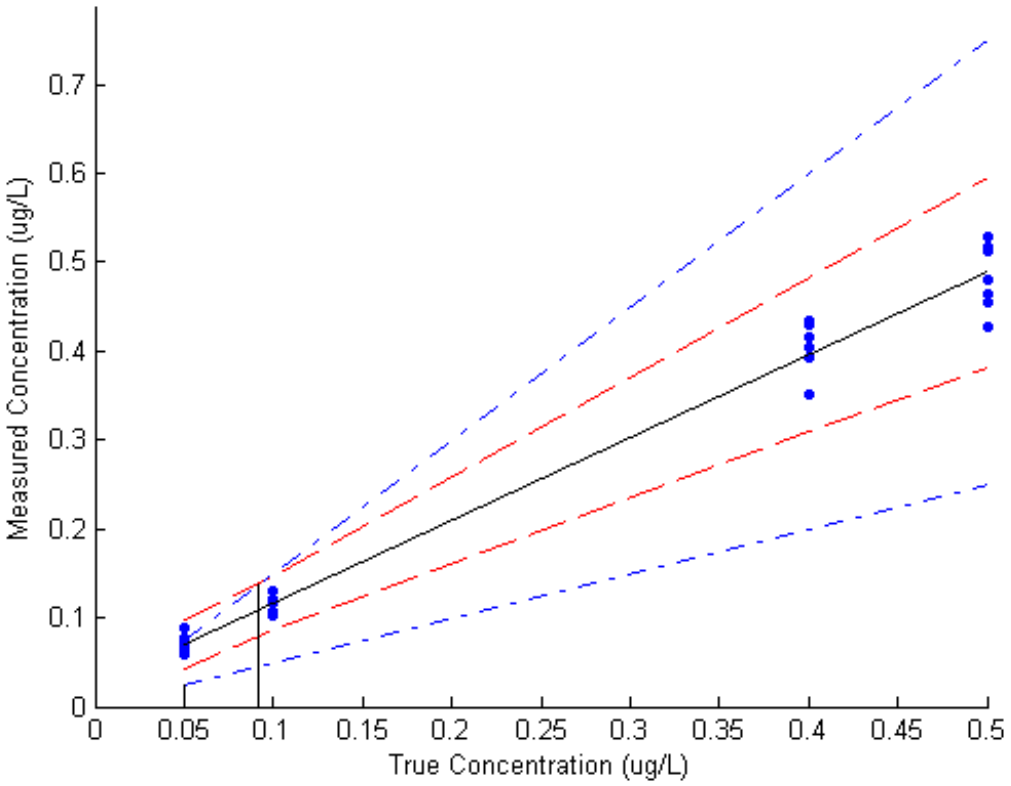
Exhibit 4: Atrazine by EPA Method 507
LCMRL Determination by Ordinary Least Squares
Outlier Excluded



$R = 0.9917$ $R^2 = 0.9836$ Adj. $R^2 = 0.9829$
$y = 0.931(x) + 0.024$
Constant Variance Test: Failed ($P>Chi^2(1) = 0.001$)
LCMRL = 0.16 µg/L

Despite exclusion of the outlier, the variance is not constant; hence, VWLS is used as shown in Exhibit 5.

Exhibit 5: Atrazine by EPA Method 507
LCMRL Determination by Variance Weighted Least Squares
Outlier Excluded



y = 0.931(x) + 0.023
LCMRL = 0.092 µg/L

The value of the LCMRL of 0.092 µg/L that is obtained by use of VWLS is lower than the value obtained by the use of OLS. This is because in the OLS model, the width of the prediction interval is highly influenced by the variance in measured values at the higher spiking concentrations. When using VWLS, the measured and spiked concentration data are weighted by dividing each value by the standard deviation of the measured concentration at its corresponding spiking concentration. As a result, the shape of the prediction interval is strongly influenced by the variance at each spiking concentration. The relatively smaller variance at the lower spiking concentrations results in a value for the LCMRL that reflects this change in variance with concentration. Criteria for determining the validity of outlier exclusion in the LCMRL determination process are presented in Section 2.2.2.

6.0 MRL VALIDATION PROCEDURE SUMMARY

Laboratories using EPA drinking water methods would be required to demonstrate that they are capable of achieving a required MRL. The procedure for validation of an MRL would be incorporated into the Initial Demonstration of Capability section of the method.

- The confidence level for the Student's *t* value and a QC interval (i.e., percent recovery) will be defined by the data users of the study. For the purposes of this paper, the two-sided confidence level for the Student's *t* value is 99% and the QC interval is 50 to 150%.

- Replicate analysis of at least seven spiked samples in reagent water are made at the MRL and are processed through the entire method procedure.

- The MRL must be contained within the range of calibration.

- A prediction interval of results (PIR), which is based on the estimated arithmetic mean of analytical results and the estimated sample standard deviation of measurement results, is determined by Equation 5 below (see also Section 7.0):

$$(5) \qquad PIR = Mean \pm s \times t_{(df,1-\alpha/2)} \times \sqrt{1+\frac{1}{n}}$$

where: *t* is the Student's *t* value with df degrees of freedom and confidence level (1-α);
s is the standard deviation of *n* replicate samples fortified at the MRL; and
n is the number of samples.

- The values needed to calculate the PIR using equation 4 are:
  i) number of replicates;
  ii) Student's *t* value with a two-sided 99% confidence level and *n* number of replicates (see Section 8.0);
  iii) the average (mean) of at least seven replicates; and
  iv) and the standard deviation of the replicate results.

- The lower and upper result limits of the PIR are converted to percent recovery of the concentration being tested. To pass criteria at a certain level, the PIR lower recovery limits cannot be lower than the lower recovery limits of the QC interval, and the PIR upper recovery limits cannot be greater than the upper recovery limits of the QC interval. When the PIR recovery limits exceed the bounds of the QC interval of recovery, the analyte fails at that concentration. If the PIR limits are contained within the bounds of the QC interval, the MRL is validated for that analyte.

During sample analysis, laboratories would need to run a daily check sample to demonstrate that, at or below the MRL for each analyte, the measured recovery is within 50% to 150%, inclusive. The results for any analyte for which 50 to 150% recovery cannot be demonstrated during the daily check would not be valid. Laboratories may elect to re-run the daily performance check sample if the performance for any analyte or analytes cannot be validated. If the performance for these analytes is validated, then the laboratory performance would be considered validated. If not, or as an alternative to analysis of a second check sample, the laboratory may re-calibrate and repeat the performance validation process for all analytes.

7.0 STATISTICAL BASIS OF MRL VALIDATION

"If a population is normally distributed with unknown mean and standard deviation, then the mean and standard deviation (s) of a random sample could be used to form a prediction interval for future observation.... If a random sample of size *n* is taken, a 100(1- α) percent prediction interval can be written as..." (Dixon and Massey, 1983)

(5) Prediction Interval = $Mean \pm s \times t_{(df,1-\alpha/2)} \times \sqrt{1+\frac{1}{n}}$

where: t is the Student's *t* value with *df* degrees of freedom and confidence level (1-α),
s is the sample standard deviation of *n* replicate samples fortified at the MRL,
n is the number of replicates.

The values needed to calculate the PIR using Equation 5 are:

i) number of replicates;
ii) Student's *t* value with a two-sided 99% confidence level and *n* number of replicates (Exhibit 6);
iii) the average (mean) of at least seven replicates; and
iv) and the sample standard deviation.

Exhibit 6: Student's *t* Values for 5 to 10 Replicates

Replicates	Degrees of freedom (df)	Student's *t* Value Confidence Level 99.0% (α/2 = 0.005)
5	4	4.604
6	5	4.032
7	6	3.707
8	7	3.499
9	8	3.355
10	9	3.25

Note: The critical t-value for a two-sided 99% confidence interval (as is used in this paper) is equivalent to the critical t-value for a one-sided 99.5% confidence interval, due to the symmetry of the t-distribution.

The factor: $s \times t_{(df,1-\alpha/2)} \times \sqrt{1+\frac{1}{n}}$

is referred to as the Half Range PIR (HR_{PIR}). For a certain number of replicates and for a certain confidence level in Student's *t*, the factor:

$$t_{(df,1-\alpha/2)} \times \sqrt{1+\frac{1}{n}}$$

is constant, and can be tabulated according to replicate number and confidence level for the Student's *t*. Exhibit 7 lists the constant factor (C) for replicate sample numbers 7 through 10 with a confidence level of 99% for Student's *t*. The HR_{PIR} is calculated by Equation 6:

(6) $$HR_{PIR} = s \times C$$

The PIR is calculated by Equation 7:

(7) $$PIR = Mean \pm HR_{PIR}$$

Exhibit 7: The Constant Factor (C) to be Multiplied by the Standard Deviation to Determine the Half Range Interval of the PIR (Student's *t* 99 % Confidence Level)

Replicates	Degrees of Freedom	Constant Factor (C) to be multiplied by the standard deviation
7	6	3.963
8	7	3.711
9	8	3.536
10	9	3.409

Note: The critical t-value for a two-sided 99% confidence interval (as is used in this paper) is equivalent to the critical t-value for a one-sided 99.5% confidence interval, due to the symmetry of the t-distribution. PIR - Prediction Interval of Results.

8.0 EXAMPLE OF VALIDATION OF LABORATORY PERFORMANCE AT OR BELOW THE MRL AND THE DAILY PERFORMANCE CHECK

Using a QC interval of recovery of 50 to 150% of the MRL and the Student's *t* confidence level of 99%, the PIR interval is calculated for aldicarb sulfoxide by entering the following values from Exhibit 10 into the PIR equation:

i) the mean of results is 0.254 μg/L;
ii) the Student's *t* value for 7 results and with a 99% two-sided confidence level is 3.707 (Exhibit 11);
iii) the standard deviation, *s*, is 0.0108; and
iv) the number of replicates, *n*, is 7.

Combining Equations 6 and 7:

$$PIR = 0.254 \pm 0.0108 \times 3.963$$

$$PIR = 0.254 \pm 0.043$$

The PIR lower result (0.254 μg/L - 0.043 μg/L = 0.211 μg/L) is converted to a percent recovery of the concentration being tested by dividing the PIR lower result by the spiked or true concentration of 0.2 μg/L and multiplying the result by 100. The percent recovery is calculated to be (0.211 μg/L / 0.2 μg/L)*(100) = 106%, which is greater than 50% and satisfies the lower limit requirement. The PIR upper result of 0.297 μg/L is converted to a percent recovery of the concentration being tested (0.297 μg/L / 0.2 μg/L)*(100) = 149%, which is less than 150%, and satisfies the upper limit requirement. The laboratory passes the MRL validation requirement for aldicarb sulfoxide at 0.20 μg/L. As seen in Exhibit 8, oxamyl and carbofuran have PIR limits that exceed the QC interval of 50 to 150% and are not validated at 0.20 μg/L in this example.

Exhibit 8: Use of PIR and QC interval with low-level data for EPA Method 531.2 (carbamates) by HPLC

	Repli-cates	True value	Mean of Results	Std Dev, *s*	PIR Half-range	PIR Lower Limit Result	PIR Lower Limit of Recovery	PIR Upper Limit Result	PIR Upper Limit of Recovery	Passes ?
		μg/L	μg/L	μg/L	μg/L	μg/L	spike	μg/L	spike	
Aldicarb sulfoxide	7	0.2	0.254	0.0108	0.0428	0.211	106%	0.297	149%	Yes
Aldicarb sulfone	7	0.2	0.204	0.0173	0.0686	0.135	67.5%	0.273	137%	Yes
Oxamyl	7	0.2	0.240	0.0168	0.0666	0.173	86.5%	0.307	154%	No
Methomyl	7	0.2	0.207	0.0205	0.0812	0.126	63.0%	0.288	144%	Yes
3-HCF	7	0.2	0.195	0.0064	0.0254	0.170	85.0%	0.220	110%	Yes

	Repli-cates	True value	Mean of Results	Std Dev, *s*	PIR Half-range	PIR Lower Limit Result	PIR Lower Limit of Recovery	PIR Upper Limit Result	PIR Upper Limit of Recovery	Passes ?
		μg/L	μg/L	μg/L	μg/L	μg/L	spike	μg/L	spike	
Aldicarb	7	0.2	0.201	0.0138	0.0547	0.146	73.0%	0.256	128%	Yes
Propoxur	7	0.2	0.203	0.0179	0.0709	0.132	66.0%	0.274	137%	Yes
Carbofuran	7	0.2	0.192	0.0341	0.1351	0.057	28.5%	0.327	164%	No
Carbaryl	7	0.2	0.180	0.0188	0.0745	0.105	52.5%	0.255	128%	Yes
1-Naphthol	7	0.2	0.210	0.0176	0.0697	0.140	70.0%	0.280	140%	Yes
Methiocarb	7	0.2	0.186	0.0183	0.0725	0.113	56.5%	0.259	130%	Yes

The confidence level for Student's *t* value is 99% and the QC interval of recovery is 50 to 150% of the level tested.

9.0 REFERENCES

American Public Health Association (APHA), American Water Works Association (AWWA), and Water Pollution Control Federation (WPCF). 1989. Standard Methods for the Examination of Water and Wastewater, 17th Ed., pp. 1-18 - 1-19 , 1030E, Method Detection Limit.

American Society for Testing Materials (ASTM) International. 2000. Standard Practice for Interlaboratory Quantitation Estimate. ASTM D 6512-00. ASTM International, 100 Barr Harbor Drive, West Conshohocken, PA 19083.

Childress, C.J.O., W.T. Foreman, B.F. Conner, and T.J. Maloney. 1999. "New Reporting Procedures Based on Long-Term Method Detection Levels and Some Considerations for Interpretations of Water-Quality Data Provided by the U.S. Geological Survey National Water Quality Laboratory," U.S. Department of the Interior, U.S. Geological Survey Open-File Report 99-193, Reston, Virginia.

Currie, L.A. 1968. "Limits for qualitative detection and quantitative determination: Application to radiochemistry," *Analytical Chemistry*, 40, p. 586-593.

Currie, L.A. 1999. "Nomenclature in evaluation of analytical methods including detection and quantification capabilities (IUPAC Recommendations 1995)," *Analytical Chimica Acta*, 391, pp. 105-126.

Dixon, W.J. and F.J. Massey, Jr. 1983. Introduction of Statistical Analysis, McGraw-Hill Book Company, Fourth Edition, p. 92.

Gibbons, R.D., D.E. Coleman, and R.F. Maddalone. 1997. "An alternate minimum level definition for analytical quantification," *Environ. Sci. Technol.*, 31, 2071-2077.

Glaser, J.A., D.L. Forest, G.D. McKee, S.A. Quave, and W.L. Budde. 1981. *Environ. Sci. Technol.*, 15, 1426, December.

Rorabacher, D.B. 1991. Statistical Treatment for Rejection of Deviant Values: Critical Values of Dixon's "Q" Parameter and Related Subrange Ratios at the 95% Confidence Level. *Analytical Chemistry*, Vol. 63, pp. 139-146.

SPSS, Inc. 1998. *SigmaPlot® 5.0 Programming Guide*. Chicago, IL.

StataCorp. 2001. *Stata Statistical Software: Release 7.0 Reference Manual*. College Station, TX: Stata Corporation.

USEPA. 1999. Revisions to the Unregulated Contaminant Monitoring Regulation for Public Water Systems; Final Rule. *Federal Register*. Vol. 64, No. 180. p. 50556, September 17, 1999.